A NEW FRAMEWORK FOR MULTI-SENSOR IMAGE FUSION

Dr. Ram_Nandan P. Singh
Naval Air Warfare Center
Aircraft Division
Patuxent River, MD 20670 - 1161

Phone : (301)342-2316

A B S T R A C T

The power of digital technology for manipulating images provides a unique, powerful tool to accomplish digital image fusion. It can facilitate fusing diverse images received from dissimilar image sensors into a composite(synthetic) image. As a result the emergence of digital image fusion technology for decomposing, comparing, mixing, and recomposing images in real time promises high potentials for applications to missile-based defense systems, and finally to a digitized battlefield.

There is a new trend in defense modernization industry to achieve a multi-mission capability for the aircraft by adding more image sensors and data links. Further, it may be noted that multiple dissimilar image sensors provide spatial, temporal, spectral, polarimetric, and other observable characteristics of an image field, and combine high intelligibility with high contrast for interesting objects or phenomena. These image characteristics greatly enhance discrimination as well as tracking of targets including small targets, and extended targets using shape, size, color, temperatures, spectral response, texture , etc., attributes of these objects in imaged scenes. Currently an image sensor processing produces mainly two types of data for target identification and tracking: (a) digitized imagery; and (b) processed data consisting of the target attributes, and the target classification result associated with its confidence level. These two types of data make to a display unit for operator evaluation. Presently these individual image sensors aboard the aircraft operate independently. As a result human operators do not get a single composite picture of a target scene. Hence there is a need to make these image sensors operate synergistically to generate a single composite picture providing a continuos, consistent image of a target area to human operators.

This research paper presents a new framework for multi-sensor image fusion in the area of digital image fusion. It enables the intelligent use of image field characteristics to generate a composite picture with high fidelity and high intelligibility. A composite data-base has been developed. This consists of sensor-level digitized data and sensor-level processed data. This new image fusion framework applies this composite data-base to digitally merge corresponding images into a composite (synthetic) picture. Two examples have been provided to demonstrate the use of this framework for image fusion applications.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-01-1998	Conference Proceedings	xx-xx-1998 to xx-xx-1998

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A New Framework for Multi-Sensor Image Fusion Unclassified	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
Singh, Ram Nandan P. ;	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS	8. PERFORMING ORGANIZATION REPORT NUMBER
Naval Air Warfare Center Aircraft Division Patuxent River, MD20670-1161	

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS	10. SPONSOR/MONITOR'S ACRONYM(S)
Director, CECOM RDEC Night Vision and Electronic Sensors Directorate, Security Team 10221 Burbeck Road Ft. Belvoir, VA22060-5806	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
See Also ADM201041, 1998 IRIS Proceedings on CD-ROM.

14. ABSTRACT
The power of digital technology for manipulating images provides a unique, powerful tool to accomplish digital image fusion. It can facilitate fusing diverse images received from dissimilar image sensors into a composite(synthetic) image. As a result the emergence of digital image fusion technology for decomposing, comparing, mixing, and recomposing images in real time promises high potentials for applications to missilebased defense systems, and finally to a digitized battlefield.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
			Public Release	14	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703767-9007 DSN 427-9007

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

1.0 INTRODUCRION

Image fusion is currently an active field of research. This research may be broadly classified into three categories: (a) fusion of multiple cues from a single image sensor, (b) fusion of images from different views with the same modality, and fusion of images from multiple modalities. In general, computer vision research based on multi-resolution techniques has been dominating this field. Now there many techniques including the wavelet transform, quadtree, and pyramid processing available for image fusion. The wavelet transform fuses transform coefficients rather than spatial image pixels, and reconstructs a fused imaged from fused transform coefficients[4,6,10]. Li, Manjunath, and Mitra [4] have applied the wavelet transform for multi-sensor image fusion, and used an integration rule that selects the larger (absolute value) of the two wavelet coefficients at each point. Wilson, Rogers, and Kabrisky[10] have performed perceptual based hyperspectral image fusion using multi-spectral analysis. They have fused the wavelet coefficients from each image using a perceptual-based weighting. Burt and Lolczynski[3] have applied pyramid processing to fuse images. Pyramid image decomposition methods include mathematical morphology and steerable pyramid decomposition. The steerable pyramid is a multiscale , multi-orientation image decomposition that uses "wavelet transform". Queiroz, Florencio, and Schaefer[7] have used a nonlinear filterbank for pyramid image coding. These methods are mathematically elegant but usually prove poor in handling complex real life spatio-temporal image fusion.

1.1 Synergistic Operation

Presently many image sensors aboard the aircraft operate independently. The integration of information across multiple human operators is nearly impossible. Hence there is a need to make all the image sensors operate synergistically that will provide a continuous, consistent picture of a scene to the decision-maker in a timely fashion. Since a multi-sensor image fusion system can take full advantage of the complementary capabilities of individual image sensors in the suite, it may produce information that cannot be obtained by viewing the sensor images separately. As a result, multi-sensor image fusion can transform incomplete, inconsistent, or imprecise data provided by individual sensors into more useful information. Past research has shown that it quite difficult to distinguish low-contrast targets from background clutter in images obtained from any single image sensor. These low-contrast targets have a weak signal-to-noise ratio (SNR). In defense applications, targets that are hard to detect in visual image can sometimes easily be noticed in a thermal image. Multi-spectral sensors can provide images that include the ultra-violet and infrared portion of the electromagnetic spectrum.

Further, the collection and production of imagery products is only valuable if it can be transmitted in time to these who need it most. Aircraft fighter pilots need target area imagery in real time; commanders require tactical reconnaissance in near real time for prompt decision-making on the battlefield. To realize these goals; the collection, fusion, and dissemination of reconnaissance and surveillance images is required in a timely fashion with accuracy and high speed. Mult-isensor image fusion will play a key role in accomplishing these goals.

1.2 Problem Definition

Multi-sensor image fusion is a complex task. As a final product, fused images may portray the specific mission needs and objectives. This research presents a general framework to intelligently fuse diverse and imprecise images from multiple dissimilar image sources into a single composite (synthetic) image in a timely fashion. The state-of-the-art technology in image processing and analysis by a single image sensor uses the frame grabber to digitize the video and sends it to the internal image analyzer of a sensor. Second, the sensor image processor derives target attributes and classification information associated with its confidence level, and outputs those processed data. Finally, these output data go to the display along with the original digitized image data for operator evaluation. In a nutshell, a single image sensor, currently, outputs two types of data:

- Processed data (target attributes, classification result, confidence level)
- Digitized image

In this research we propose a new scheme for fusing images received from two or more image sensors. A new scheme for fusing imagery from multiple sensors is proposed. Here we store processed data in a database called "knowledge-base", and digitized image data in " digital imagery". This scheme gives rise to several new problems. The first and foremost is image registration [2]. Images received from two separate sensors need to be registered. Broadly speaking there are two types of registration: temporal registration and spatial registration. These align the imagery data both in space and time. The spatial registration of the images corrects for relative translation shifts as well as geometrical and intensity distortions of each image [1]. Given two or more pictures of the same object by different sensors and their registration, we can determine the characteristics of each pixel with respect to all of the sensors. In this paper we assume that the images to be combined are already perfectly registered.

In this research, we present an image fusion framework that employs fuzzy pattern combination to fuse input images. Section 1 presents the theoretical foundation to realize this framework, and Section 3 provides an image fusion methodology . Results are presented in Section 4.

2.0 BASIC FRAMEWORK

The availability of multi-spectral image sensors along with multiple dissimilar image sensors aboard an airborne platform has created a need for a basic image fusion framework that can fuse images from dissimilar sensors or sequential images from the same sensors in real time. Since an enormous amount of data is needed to describe a single image, an image fusion framework calls for a generic methodology based on digitization and intelligent fusion. It has been well-said that an image is worth thousand words, and a composite (synthetic) image is worth thousand images. This framework provides the foundation for a seamless flow of information among all tactical, strategic, and sustaining base systems. Finally, this helps realize a digitized battlefield which is the key to a more efficient and effective fighting force in an arena of declining service budgets and reduced force structure.

In this section we develop a theoretical framework for understanding deep structure of an image with a view to developing a sound basis for intelligent image fusion in the sequel. The

simplest method for fusing images is accomplished by computing their average. Although the features from each image, to a varying degree, are present in a fused image, the contrast of the original feature can be significantly reduced and blurred. As a result, this can render the fused image useless for practical applications. Another simple method ids to use the maximum values on intensity to generate a fused image. Even though this method is analytically tractable, it can miss the minimum values (which are critical to the mission) submerged in the processing. To overcome this deficiency, we can use max-min techniques for fusion. But preserving the shadows to provide various textures and patterns are critical to a final product.

2.1 Basic Structures of Images

An image could represent luminance of objects in a scene (picture by an ordinary camera), absorption characteristics of body tissues(X-ray imaging), the radar cross-section (radar imaging), temperature profile of a region(infrared imaging), and many more. In practice these images have often imprecise boundaries and broad description of details. J. K. Hawkin echoed this in his statement: "In reality, in our normal visual environments, no two objects are exactly alike. An organism that wants to survive or a device that has to act intelligently in such an environment, must be able to disregard variations which are unimportant at a particular instant. Only then can the visual environments become describable in terms of rather loosely defined sets of objects, sets of actions which the object is capable of or which it is useful for". This is the primary guideline that provides a basis for a meaningful image fusion.

The basic structure of a human visual perception model provides a basis for understanding the image fusion processing of diverse images by human beings. The brain samples and represents the optic array at many resolution scales simultaneously. As a result, the visual system represents retina images at all levels of resolution simultaneously, and deals with contrast, spatial frequencies, and color. This reveals the hierarchical architecture of this model with links between the different levels of resolution. This structure fits into existing theories of the visual system as a continuous stack of homogeneous layers characterized by iterated local processing schemes which, in turn, point to pyramid processing. The three different formulations that define the deep structure of an image are the diffusion of an image characterized by a parabolic linear partial differential equation of the second order, the convolution of an image with a family of Gaussian point spread functions, and the iterated blurrings of an image (which asymptotically leads to diffusion) in an apparent ad hoc fashion.

2.2 Image Field Characteristics

The development of new imaging sensors has opened up new areas for image processing. These include the intelligent use of spatial, spectral, polarimetric, and temporal characteristics of an image field to "synthesize" images, which combine high intelligibility with high contrast for interesting objects of phenomena. Since we can measure the polarization state, the phase, and the amplitude; we can take advantage of these three attributes to-gather for object recognition by machines, whereas the human eye only sees amplitudes.

In practice, images are very often have imprecise boundaries and broad descriptions of details. As a result, fusing images into a composite image in a timely fashion is a problem intrinsically incapable of precise mathematical formulation. We have to deal with imperfect

measurements. These measurements may be both imprecise and uncertain. In this context, Albert Einstein had observed," So far as the laws of mathematics refer to reality, they not certain; and so far as they are certain, they do not refer to reality". L. A. Zadeh has reinforced this idea and remarked: "As the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a thresh hold is reached beyond which precision and significance (or relevance) become mutually exclusive characteristics " [1]. In a nutshell, high precision is incompatible with high complexity. However, human operators can tolerate some degree of imprecision in detail. As a result, conceptual structure of the theory of fuzzy sets and fuzzy logic may well provide a natural setting for the formulation and approximate solution of complex, nonlinear problems in image fusion.

2.3 Development of Image Fusion Framework

There are two essential requirements for image fusion: (1) pattern conservation: important details of the component images must be preserved in the composite image; and (2) spurious elements avoidance: it must not introduce any new pattern elements or artifacts that could interfere with subsequent image analysis and reconstruction. Because spatial information is essential for reconstructing a composite image, it cannot be sacrificed while decomposing an image into a set of primitive patterns. This is the key to putting up a successful image fusion scheme in practice. Keeping that in mind, an intelligent image fusion theory has been developed based on diffusing images with respect to contrast, spatial frequencies, and color into a set of diffused images; and molding the diffused images into a composite image. This involves generating a set of diffused images from a given set of candidate images, combining them into a composite set of diffused images, and , finally, compacting them into a composite image. Based on this theory, the design of an image fusion paradigm involves decomposing a set of fuzzy primitives, defuzzifyimg each composite primitive, and reconstructing a composite image from the set of composite primitives.

In a generic image fusion architecture for intelligent image fusion, each image sensor outputs two type of data: (a) raw digitized image data and (b) processed image data (target attributes, classification result, and confidence level). All these data go to a composite knowledge base consisting of digitized and processed data. In addition it shall also contain domain expertise, sensor characteristics, and fuzzy rule-base. Clearly this composite knowledge-base contains both numeric as well as linguistic information. The data flow and the processing mechanism to process these data at different stages of image fusion processing involves various transformations and, combinations. Digitized data from each input image undergo a decomposition transformation that decomposes the input image into a number of pyramid levels. For the purpose of image fusion, each resolution level may be considered a single or multiple pattern elements. Next, it goes through a fuzzy system paradigm that performs fuzzification, combination, and defuzzification of the composite set of primitives. Finally, an inverse pyramid transformation is applied to reconstruct a composite image.

3. IMPLEMENTATION

Primary difficulties in fusing diverse images arise from the fact that scene information gathered using various image sensors (ISAR/SAR radar, IRDS, EO, etc., ..) possess imprecise boundaries and varying resolutions derived from uncertain, dynamic environments. In addition, an image understanding task becomes difficult because of complex shapes, uneven illuminations, shadows, and complex textures. As a result, all these factors make it difficult to design general algorithms to solve real life image fusion problems.

3.1 Generic Pyramid-Based Image Fusion

A pyramid representation provides a structure to merge primitive elements/attributes on a local level, and subsequently on a global (I.e. object) basis. A generic image fusion method is a four-stage process to accomplish this goal. It consist of the following:

* Decompose each source image into primitive elements
* Determine suitable criteria to merge primitive pattern at each pyramid level
* Combine source pyramids to form a composite pyramid
* Regenerate the composite image through an application the inverse pyramid transformation

This methodology requires crisp images for processing.

3.2 Composite Data Base Construction

In practice there is no crisp hierarchy of layers in a pyramid. There is always some degree of overlap between adjacent layers with gradual transitions into one another. Thus each layer can be modeled with one or more fuzzy membership functions. Sometimes each layer can represent separate attributes. Using fuzzy rules these attributes/ primitives can be combined. It consists of the following possessing:

* Decompose each source image into a set of primitive pattern elements.
* Fuzzify attributes and/ or pattern elements
* Develop fuzzy rules to merge fuzzy pattern elements at each pyramid level
* Apply these rules to merge the set of fuzzy primitives to form a single set of fuzzy primitives for the composite image at each pyramid level
* Defuzzify each element of the composite pyramid. This will produce a composite pyramid of crisp elements.
- Reconstruct a composite image from the composite pyramid using an inverse pyramid transform

3.3 Levels of Image Fusion

- There are three important levels of image fusion:
- Element level fusion - Employs pixels or primitive image patterns. This is low-level fusion which uses basic information
- Attribute level fusion - Intermediate level fusion, which uses derived information from pixels or image primitives.
- Decision level fusion - High-level fusion which uses merging rules

It may be noted that single level fusion ignores some of the information, which is available at other levels. Multi-level fusion uses most of the available information across the three levels of fusion.

3.4 Basis for Merging Rules

Detecting of objects in imagery can be accomplished by the analysis of amplitude, spatial, spectral, and temporal characteristics of two-dimensional images. Rules for merging primitive pattern elements/attributes may be based on the following structures:

- Hierarchical Reasoning - Uses image structures to frame merging rules. It implies a ranking or ordering process. Human perception involves a hierarchical reasoning process that relies on both bottom-up and top-down reasoning. Situation awareness often demands both bottom-up and top-down reasoning.
- Spatial Reasoning - Exploits spatial frequencies and relationships between elements of different images. It implies three-dimensional character of the physical world.
- Temporal Reasoning - Deals with motion images. It involves with dynamic and evolving situations in space and time.

Using the fuzzy rule, the confidence levels associated with different target IDs determined independently by individuals sensors can be combined into a composite pattern confidence event.

4.0 RESULTS

There are four critical components of multi-sensor image fusion processing: (a) image registration, (b) image normalization, (c) image correlation, and (d) image fusion (or update). For the purpose of demonstrating, the intelligent image fusion theory, it is assumed that all data are properly registered and normalized. An example is provided for illustration.

Given two source images, obtain a composite picture. The following steps are involved:

Step # 1: Perform convolution with a source image using a Gaussian function as a weighting function (Eqn. (1)). This will generate Gaussian pyramid#1 for the input image#1.
Step # 2: Derive Laplacian pyramid#1 from Gaussian pyramid#1 following the methodology as outlined in Section 2.2.
Step # 3: Repeat the above two steps for source image#2. This will generate Gaussian pyramid#2 and Laplacian pyramid#2.
Step #4: Fuzzify each level in both Laplacian pyramids.

Step #5: Develop a set of fuzzy rules to determine the degree of correlation between two images.
Step #6: Apply fuzzy logic to compute the correlation values for each level.
Step #7: Merge the primitives into a composite set of primitives. Construct a fuzzy composite Laplacian pyramid.

Step # 8 Defuzzify each fuzzy primitive in the fuzzy Laplacian pyramid. This will generate a crisp composite Laplacian pyramid.
Step #9: Regenerate a composite image from the composite Laplacian pyramid.

To fuzzify each level in a crisp Laplacian pyramid, one can model it with a single or multiple fuzzy membership functions [9,12]. Two key fuzzy variables for deriving fuzzy rules are: intensity differences and intensity gradients. Fuzzy rules have been developed to provide various combinations of these two fuzzy variables with varying degrees of fuzziness[2,9].

4.1 Example

The two input images of the same object are selected. The input #1 is shown on the left-bottom,   of Figure -1 and the input #2 on the right bottom. Figure 1 shows that the input image #1 has more details whereas the input image #2 is more blurred. Figure - 1 presents their Gaussian pyramids in a compact form. Figure – 2, depicts the Laplacian pyramid1 derived from the Gaussian pyramid1, and similarly Figure-3 shows the Laplacian pyramid2 for the Gaussian pyramid2. Figure – 4, depicts the composite Laplacian pyramid after the two Laplacian pyramid1 and Laplacian pyramid2 have been combined. Figure -5 presents the composite image obtained from the composite Laplacian pyramid.

5.0 CONCLUSIONS

An intelligent image fusion paradigm based on the theoretical frame-work developed has been implemented. Fuzzy system techniques provide a powerful methodology to merge image primitives in real situations. It has been demonstrated that the composite image contain more details about the object than any single input image. Potential applications for defense include direct combat target identification in real time and underwater mines detection, classification, and identification.

6.0 REFERENCES

1. Brown, L., " A Survey of Image Registration Techniques", ACM COMPUT. Survey, 24, pp. 325-376, 1992

2. Burt, P. J., "The Pyramid as a Structure for Efficient Computation," in Multi-Resolution Image Processing and Analysis, A. Rosenfield, ed., Springer-Verlag (1984)

3. Burt, P. J. & Lolczynski, R. J. , " Enhanced Image Capture through Fusion", in Proc. Of the Fourth International Conference on Computer Vision, Berlin, Germany, May 1993, pp. 909-960.

4. Li, H; Manjunath , B. S.; and Mitra, S. K.; "Multi-sensor Image Fusion Using the Wavelet Transform", Graphics Models and Image Processing, Vol. 57, No. 3, May 1995, pp. 235 - 245.

5. Luo, R., and Kay, M., "Data Fusion and Sensor registration: State-of-the-art in 1990s" in Data Fusion in Robotics and Machine Intelligence, (Abidi, M. and Gonzaloz, R., Eds.), pp. 7-136, Academic Press, San Diego, 1992.

6. Mallat, S.G., "A Survey of Multiresolution Signal Decomposition: Wavelet Representation", IEEE Trans. On Pattern Analysis & Machine Intelligence, PAMI -11, July 1989, pp.674-693

7. Queiroz, R.L. de; Florencio, D. A. F., and Schaefer, R. W. , " Nonexpensive Pyramid for Image Coding Using a Nonlinear Filterbank", IEEE Trans. On Image Processing, Vol. 7 No.2, Feb. 1998, pp.246-252

8. Toet, A. "Hierarchical Image Fusion", Machine Vision Applications, March 1990, pp. 1-11

9. Singh, R-N. P. and Bailey, W. H. , "Fuzzy Logic Applications to Multisensor - Multitarget Correlation", IEEE Trans. on Aerospace and Electronic Systems ,Vol. 33, No. 3, July 1997

10. Wilson, A. T.; Rogers, S. K., Kibrinski, M., ""Perceptual-Based Image Fusion for Hyperspectral Data", IEEE Trans. On Geoscience & Remote Sensing, Vol. 35, No. 4, July 1997, pp. 1007-1017

11. Zadeh, L. A., "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. Systems Man & Cybernetics, 1 28-44 (1973)

12. Zadeh, L. A. , "Fuzzy Sets and Information Granularity," in M.M. Gupta, R.K. Ragade, R. R. Yager, ed., Advances in Fuzzy Set Theory and Application, North-Holland (1979)
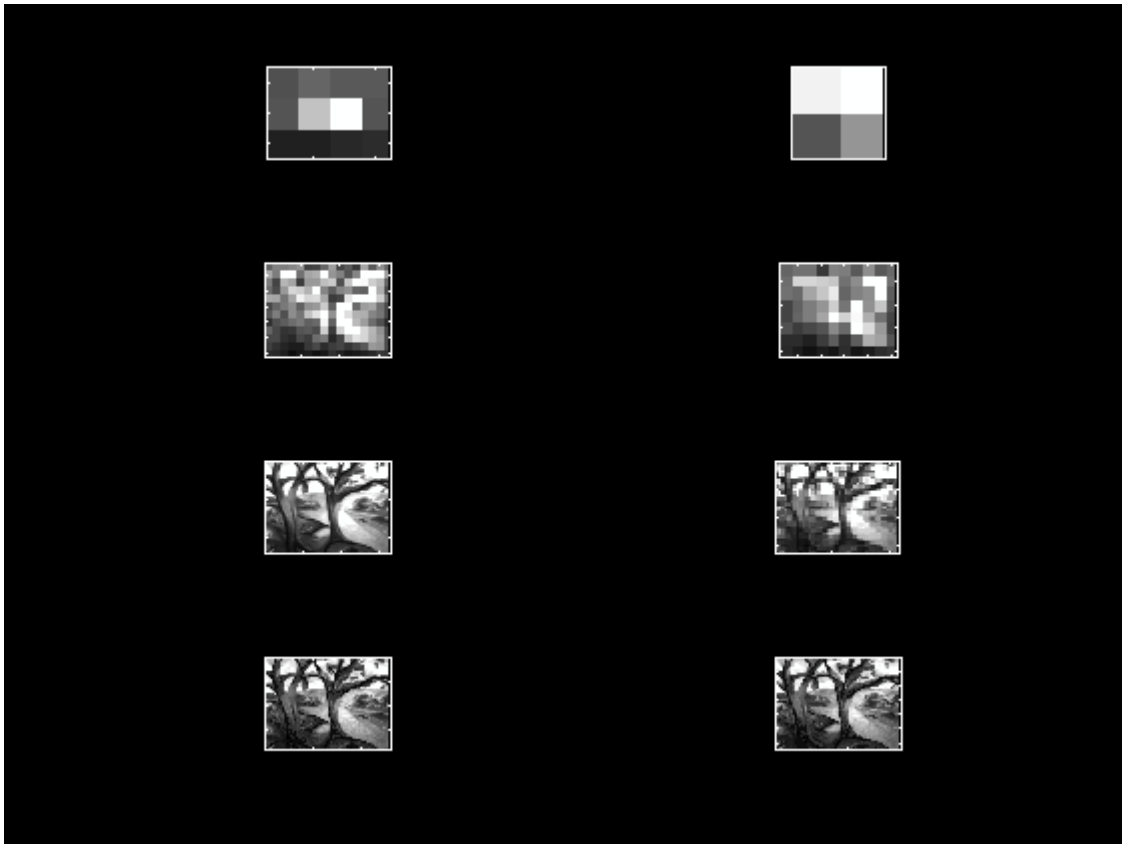
GAUSSIAN   PYRAMID1   GAUSSIAN   PYRAMID2
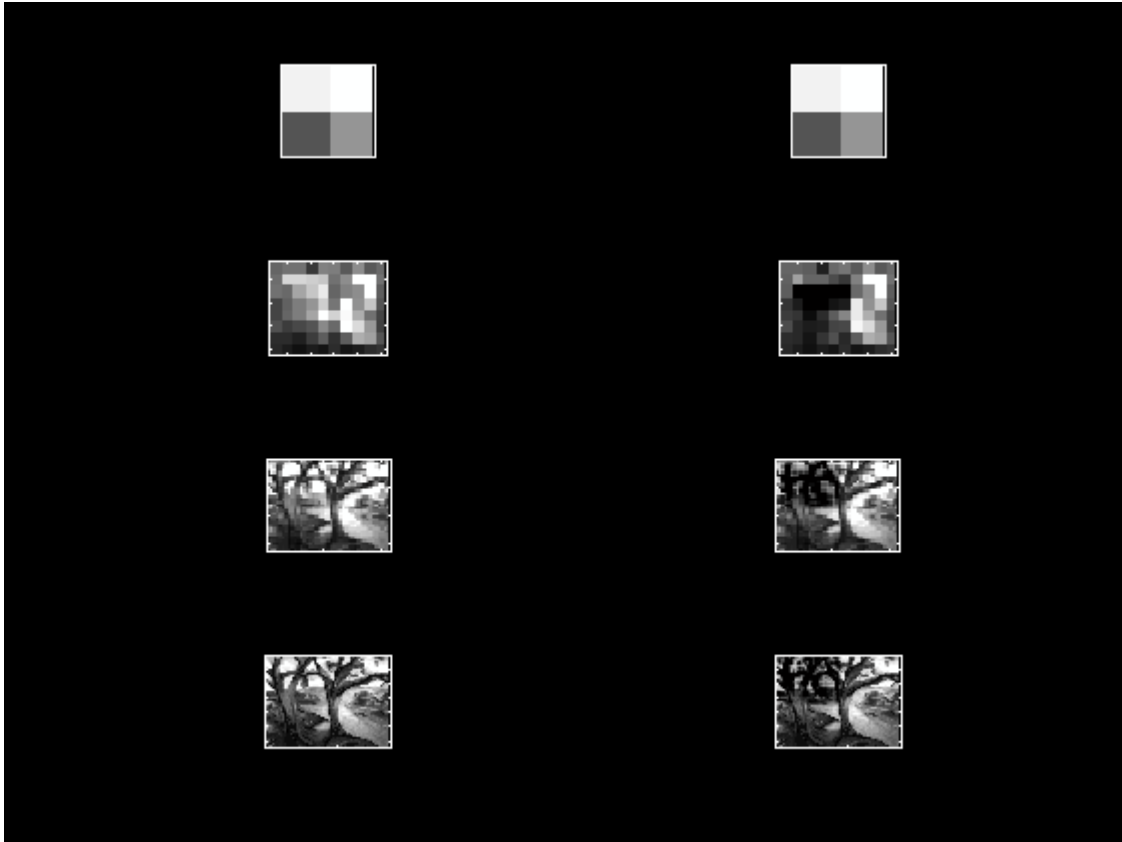
FIGURE - 1

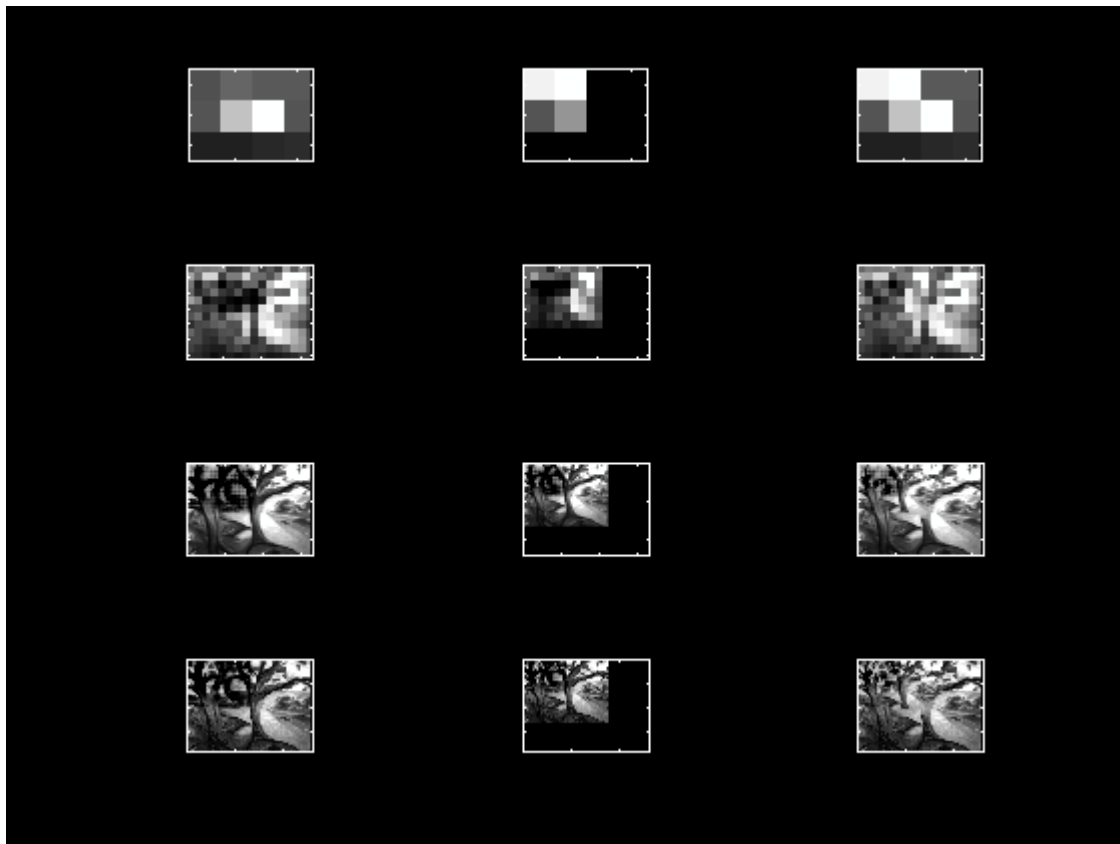GAUSSIAN PYRAMID1   LAPLACIAN PYRAMID1

FIGURE - 2

GAUSSIAN PYRAMID2   LAPLACIAN PYRAMID2

FIGURE - 3

LAPLACIAN
PYRAMID1

LAPLACIAN
PYRAMID2

COMPOSITE
LAPLACIAN

FIGURE - 4

COMPOSITE LAPLACIAN   RECONSTRUCTED IMAGE
        PYRAMID

FIGURE- 5